



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Recombination changes at the boundaries of fully and partially sex-linked regions between closely related *Silene* species pairs

Citation for published version:

Charlesworth, D, Campos, JL, Qiu, S, Guirao-Rico, S & Bergero, R 2016, 'Recombination changes at the boundaries of fully and partially sex-linked regions between closely related *Silene* species pairs', *Heredity*, vol. 118, pp. 395-403. <https://doi.org/10.1038/hdy.2016.113>

Digital Object Identifier (DOI):

[10.1038/hdy.2016.113](https://doi.org/10.1038/hdy.2016.113)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Heredity

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Recombination changes at the boundaries of fully and partially sex-linked regions
between closely related *Silene* species pairs.

Authors

JL Campos, S Qiu¹, S Guirao-Rico², R Bergero, D Charlesworth

Institute of Evolutionary Biology, School of Biological Sciences, University of
Edinburgh, Edinburgh, UK

¹ Current address: Department of Animal and Plant Sciences, University of Sheffield,
Sheffield, UK

² Current address: Department of Animal Genetics, Center for Research in
Agricultural Genomics (CRAG) CSIC-IRTA-UAB-UB, Edifici CRAG,
Campus Universitat Autònoma Barcelona, Bellaterra, Spain

Corresponding author: D Charlesworth

Institute of Evolutionary Biology, School of Biological Sciences, University of
Edinburgh, Edinburgh, UK

Email: Deborah.Charlesworth@ed.ac.uk

Fax: ++44 131 650 5751

Keywords: Sex chromosomes, recombination suppression, pseudoautosomal region,
partial sex-linkage

Running title: Evolution of sex-linkage in *Silene*

Data Archival Location: Dryad

Word count for main text: 6874, including all main text, acknowledgements etc, but
excluding references

Abstract

The establishment of a region of suppressed recombination is a critical change during sex chromosome evolution, leading to such properties as Y (and W) chromosome genetic degeneration, accumulation of repetitive sequences, and heteromorphism. Although chromosome inversions can cause large regions to have suppressed recombination, and inversions are sometimes involved in sex chromosome evolution, gradual expansion of the non-recombining region could potentially sometimes occur. We here test whether closer linkage has recently evolved between the sex-determining region and several genes that are partially sex-linked in *Silene latifolia*, using *S. dioica*, a closely related dioecious plants whose XY sex chromosome system is inherited from a common ancestor. The *S. latifolia* pseudoautosomal region (PAR) includes several genes extremely closely linked to the fully Y-linked region. These genes were added to an ancestral PAR of the sex chromosome pair in two distinct events probably involving translocations of autosomal genome regions causing multiple genes to become partially sex-linked. Close linkage with the PAR boundary must have evolved since these additions, because some genes added in both events now show almost completely sex-linkage in *S. latifolia*. We compared diversity patterns of five such *S. latifolia* PAR boundary genes with their orthologues in *S. dioica*, including all three regions of the PAR (one gene that was in the ancestral PAR, and two from each of the added regions). The results suggest recent recombination suppression in *S. latifolia*, since its split from *S. dioica*.

INTRODUCTION

Sex chromosome evolution involves the establishment of a region of suppressed recombination between a chromosome pair that carries a sex-determining locus (reviewed by Bull 1983). If a non-recombining region persists, it may lead to differentiation of the Y from the homologous X chromosome (or between Z and W chromosomes in ZW systems). Regions where recombination still occurs between the X and the Y are called pseudo-autosomal regions (abbreviated to PAR). In several species, the region of suppressed recombination has expanded in successive events, each time shrinking the PAR and forming a new fully sex-linked region whose alleles in the non-recombining Y or W chromosome then start diverging from those of the homologous X or Z chromosome. Consequently, genes in X chromosome regions physically closest to the PAR have lower sequence divergence than ones distant from the PAR. These regions of different Y-X divergence are termed "evolutionary strata", and were first noticed in human sex chromosome sequences, and later found in all mammals (Cortez et al. 2014), and in the ZW chromosomes of birds, although some paleognathous birds have retained extensive PARs (reviewed in Zhou et al. 2014). Two plants that have separate sexes (dioecious plants) and sex chromosomes with XY males and XX females, *S. latifolia* and *Carica papaya*, also have strata (Bergero et al. 2007; Wang et al. 2012). In papaya, two distinct strata appear to have evolved by fixation of inversions in the Y-linked region (Wang et al. 2012), but in *S. latifolia* Y-X divergence may increase without discontinuity with distance from the PAR (Chibalina and Filatov 2011).

The reasons for the repeated formation of new evolutionary strata, or for the possible gradual extension of sex chromosomes' non-recombining regions, are not yet fully understood, and the mechanisms involved are not completely known in most species, with new mysteries emerging about the PAR boundary even in well-studied species like humans (Cotter et al. 2016). One plausible reason/cause for the situation driving changes in the PAR boundary (and creation of new strata) involves sexually antagonistic mutations (advantageous in one sex but disadvantageous in the other, often abbreviated to SA). SA mutations in genes closely linked to the fully sex-linked region may establish polymorphisms, leading to selection for closer linkage. For example, closer linkage between male-benefit SA alleles and the Y-linked region

restricts these alleles to males, and avoids harming the females (Bull 1983; Rice 1987; Jordan and Charlesworth 2012).

Recently evolved sex chromosomes, such as those in some plants, offer good systems for studying how partially sex-linked regions evolve full sex linkage. In the plant *S. latifolia*, recombination suppression between the fully Y- and X-linked regions was initiated around 5-10 MY ago, but at least one region evolved suppressed recombination subsequently, forming a younger fully Y-linked male-specific, or MSY, region (Bergero et al. 2007; Chibalina and Filatov 2011). By comparing the genetic map of the *S. latifolia* X chromosome with mapping results in the related species *S. vulgaris*, which does not have sex chromosomes, the *S. latifolia* PAR was inferred to have been formed by independent additions of two genomic regions to an ancestral PAR, through translocations from other chromosomes, as shown in Figure 1 (Bergero et al. 2013; Qiu et al. 2016). Several genes located near the boundary with the fully sex-linked region recombine so rarely that some variants are found only in males (Qiu et al. 2016); among these PAR boundary genes, some were added in one addition event, and some in the other (Figure 1).

Figure 1 here

Translocations that add new regions onto a PAR, as in *S. latifolia* and its close relatives, are particularly interesting in relation to the evolution of suppressed recombination, as crossing over in the ancestral PAR should not be affected by an addition, and should continue after the addition event, generating recombinants between the MSY and the added regions. Therefore, if such recombination does not occur, it was probably suppressed after the rearrangement occurred. Such an addition occurred in the ancestor of Eutherian mammals, and recombination was indeed subsequently suppressed, moving the PAR-MSY boundary far from its pre-addition location (reviewed in Cortez et al. 2014). However, few such cases have been studied. The genes near the *S. latifolia* PAR boundary represent a much more recent translocation situation than that in mammals, and are ideal for testing whether the PAR-MSY boundary has remained in the same location. Close linkage observed between the *S. latifolia* MSY and genes from both addition events (Qiu et

al. 2016) suggests recombination suppression after the translocations occurred (other genes added in both translocations, have remained loosely linked to the MSY boundary). However, although the translocations should not directly suppress recombination in the PAR boundary region, it is important to test this alternative.

We therefore examined species closely related to *S. latifolia*, mainly, but not exclusively, the closest relative, *S. dioica*. The *S. dioica* XY chromosomes are homologous with those of *S. latifolia* (Nicolas et al. 2005), and have indistinguishable morphology and arm ratios (Grabowska-Joachimciak and Joachimciak 2002). The two species hybridise readily, producing fertile progeny, consistent with non-rearranged chromosomes, and there is evidence for ongoing gene flow (Muir et al. 2012; Hu and Filatov 2015). If the translocation events outlined above occurred in a common ancestor of these species and directly suppressed recombination between the added regions and the MSY region, the PAR boundary genes should also show close linkage to the MSY in *S. dioica*. If, however, closer linkage with the Y-linked region has evolved in *S. latifolia*, linkage should be looser, or absent, in *S. dioica*.

Three other closely related *Silene* species, *S. diclinis*, *S. marizii* and *S. heuffellii*, are also dioecious. Reliable inference of the order in which they split will require large numbers of gene sequences, which are not currently available other than for *S. latifolia* (Rautenberg et al. 2010). However, *S. latifolia* forms hybrids more readily with *S. dioica* than *S. diclinis* (Prentice 1978), suggesting that *S. diclinis* is an outgroup to the two other species. Genes that are fully sex-linked in *S. latifolia* are also fully sex-linked in *S. dioica* and *S. diclinis* (Laporte et al. 2005; Nicolas et al. 2005; Kaiser et al. 2009; Muir et al. 2012), supporting the view that the sex chromosomes of these species are homologous. *S. marizii* and *S. heuffellii* have been less studied, but, consistent with the genetic evidence just outlined, the sex chromosomes of *S. marizii* are similar in morphology to those of *S. latifolia* and *S. dioica*, whereas *S. diclinis* has undergone a Y-autosome reciprocal translocation (Howell et al. 2009). As explained below, we demonstrated that one gene, *SLCyp* in the younger *S. latifolia* sex chromosome stratum, with an estimated K_s value between the X- and Y-linked sequences of 0.067 (Bergero et al. 2007), also has variants found only in males of *S. dioica*, *S. diclinis* and *S. marizii*, indicating that this recently evolved stratum is sex-

linked in all of them. However, *S. latifolia* PAR genes have not yet been mapped in the outgroup species.

We here study outgroups to ask (i) whether the *S. latifolia* PAR boundary genes are as well closely linked to the MSY in the outgroup species, and to ask the related question (ii) whether results suggest that the translocations that caused these genes to become partially sex-linked in *S. latifolia* directly caused restricted recombination, versus recombination being suppressed subsequently. Rather than genetically mapping these genes, we used a highly sensitive population genetic approach, testing for associations between SNPs in the PAR boundary loci and the male-determining region, using the subdivision measure K_{ST} between males and females. K_{ST} is the average F_{ST} per site, computed from sequence data (Hudson et al. 1992), and reflects linkage disequilibrium due to population subdivision (Charlesworth et al. 1997), such as Y-X differentiation resulting from absent or very rare recombination. F_{ST} has already been proposed as the best way to test for associations between alleles of partially sex-linked genes and a fully sex-linked locus (Qiu et al. 2013; Kirkpatrick and Guerrero 2014; Qiu et al. 2016). Importantly, this can potentially detect recombination even if it is too rare to be detectable in families, as is the case in *S. latifolia* for the PAR boundary loci studied here (Qiu et al. 2016). Our analysis suggests recently decreased recombination in *S. latifolia*. Did the translocations directly cause recombination suppression? Based on evidence of partial sex linkage in *S. dioica* for the two genes we studied in the first translocation, subsequent recombination suppression is implied between the *S. latifolia* MSY and the added region. The second translocation could, however, have occurred in *S. latifolia* during the very short evolutionary time separating this species from *S. dioica*, so it remains possible that this event directly suppressed recombination (though we present arguments that this is unlikely).

METHODS

Genes and plant samples

In *S. latifolia*, no recombinants were detected between the genes studied here and the MSY region in male meiosis, and, in natural populations of *S. latifolia*, they all show marked sex differences in allele frequencies, including some variants found

only in males, indicating very close linkage to the MSY (Qiu et al. 2016). Gene *cs3597* is part of a putative ancestral PAR, as, in *S. vulgaris*, it maps to the same linkage group (SvLG12) as the genes that are fully sex-linked in *S. latifolia* (Qiu et al. 2016). The four other *S. latifolia* PAR-MSY boundary region genes studied here map to two other linkage *S. vulgaris* groups. Specifically, the sequences of genes *E559* and *E521* were added along with 4 other genes that map to linkage group SvLGSmall and map far from the MSY in *S. latifolia*, while, of the 7 *S. latifolia* PAR genes that map to SvLG9, only *cs935* and *E523* are in the boundary region (Qiu et al. 2016). We also sequenced a gene, *cs4991*, located slightly more distal to these PAR boundary genes. No recombinants were seen in males in the family in which *cs4991* was mapped; however, it maps very close to gene *E352* in the X chromosome, which yielded many recombinants (6/58) in male meiosis of another *S. latifolia* family, so the recombination frequency in the *S. latifolia* population as a whole could be several percent (Qiu et al. 2016). The orthologue of *cs4991* has not been mapped in *S. vulgaris*, but its map location in *S. latifolia* suggests that it was added to the PAR along with other SvLG9 genes, after the SvLGSmall genes were added. All these genes segregate as single-copy loci in at least one *S. latifolia* full-sib family, with two alleles in both sexes (Bergero et al. 2013; Qiu et al. 2016).

We sequenced orthologues of the *S. latifolia* PAR genes in *S. dioica* males and females, using plants grown from seeds collected from 11 different locations distributed across Europe (see Supplementary Table S1 and Figure S1, which also shows the localities from which the *S. latifolia* samples were collected; these are described in detail in Qiu, Bergero and Charlesworth, 2013). The samples from each species were collected from many different populations for two reasons. First, this ensures that the results are representative of the species generally. Second, our goal was to test for associations between sequence variants and the sex-determining region. Such a “scattered sample”, with few individuals per population, minimizes false inferences caused by random associations within populations (Städler et al. 2009).

Plants were grown in Edinburgh and sexed once they flowered. The sexes were also confirmed by PCR amplification of intron 2 of *SlCyp*, a gene in the younger of the two evolutionary strata in this species (Bergero et al. 2007). The sexes assigned by

this marker agreed perfectly with those observed at flowering, indicating that none of our males carries a recombinant genotype for the sex chromosomes.

In total, we obtained sequences for most of the genes from 12 alleles from *S. dioica* females, and 20 from males, so as to represent alleles associated with the Y as well as the X chromosome. The *S. latifolia* samples with which we compared these new results included 38-44 alleles sequenced from females and 40-42 from males (Qiu et al. 2016). Smaller samples of *S. marizii* and *S. diclinis* (see Supplementary Table S1) were also studied for two genes, *E559* and *E523*, one from each addition event.

DNA extraction, PCR reactions, and cloning

Genomic DNA for sequencing was extracted from leaves using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. PCR amplifications were then performed with primers given in Qiu et al. (2016), using Phire Hot-Start DNA Polymerase (Thermo Fisher Scientific, Paisley, UK) in a Finnzymes' Piko cycler, with the following conditions: 1 cycle of initial denaturation at 98° for 30 sec, 10 cycles of DNA denaturation at 98° for 5 sec, primer annealing varying from 60 to 70° for 5 sec, and DNA amplification at 72° for 30 sec, 25 cycles at 98° for 5 sec, 60° for 5 sec, 72° for 60 sec, and finally 1 cycle at 72° for 5 min. The PCR products were cleaned with ExoSAP-IT (Affymetrix, High Wycombe, UK) and Sanger sequenced on an ABI 3730 capillary sequencer (Applied Biosystems, Warrington, UK). Single nucleotide polymorphisms (SNPs) were examined after direct sequencing. DNAs producing PCR amplicons with heterozygous indels (generally in intron regions) were re-amplified with the Phusion proof-reading DNA polymerase (Thermo Fisher Scientific, Paisley, UK) using the same PCR conditions as above, and cloned into the vector pJET 1.2/blunt (Thermo Fisher Scientific, Paisley, UK) before sequencing, six to eight colonies were screened per amplicon. The resulting sequences were aligned in Sequencher 4.8 (Gene Codes, Ann Arbor, MI; <http://www.genecodes.com>), including sequences of the orthologous genes from *S. latifolia* and *S. vulgaris*, and manually adjusted using Se-al v. 2.0 (<http://tree.bio.ed.ac.uk/software/seal/>). Alignments including sequences from all species studied here have been deposited in Dryad under accession numbers [TO BE ADDED].

Sequence analyses

We compared variant frequencies of alleles, and frequencies of heterozygotes in males and females, using Chi-square tests. The polymorphism analyses, including estimates of nucleotide diversity and Tajima's D (Tajima 1989), and divergence estimates between the *S. latifolia* and *S. dioica* sequences were done using DnaSP v.5.00.06 (Librado and Rozas 2009). Estimates of K_{ST} and tests for subdivision under the null hypothesis of no subdivision were done using 1000 permutations of K^*_{ST} and another subdivision measure, S_{nn} (Hudson et al. 1992) implemented in DnaSP. Significance of K_{ST} was tested using the K^*_{ST} statistic, rather than with K_{ST} , because K^*_{ST} has good power for small samples (Hudson et al. 1992). To obtain error bars for the K_{ST} values shown in Table 1 and Supplementary Figure S4, we resampled individuals 40 times from each species with replacement, maintaining the same sample sizes of males and females for each gene, and re-estimated K_{ST} using these new alignments for each gene and species.

Our samples from *S. dioica* are smaller than from *S. latifolia*, which will reduce the power of our tests of significance with *S. dioica* samples. To compare associations between variants and the fully sex-linked region, we therefore sub-sampled from *S. latifolia*, as follows. For each gene, we constructed samples with the same numbers of males and females as in the *S. dioica* sample, by randomly sampling individuals separately from the two sexes to form a reduced sized *S. latifolia* sequence sample. We estimated the means of Tajima's D from 1 000 such sub-samples, using the batch mode of DnaSP. However, as the batch mode does not estimate K_{ST} or perform tests of significance of subdivision, we used 40 sub-sampled alignments to assess the effect of sample size on this measure of subdivision, and its significance, in the manner just described.

RESULTS

Subdivision between *S. latifolia* and *S. dioica*

Because there is ongoing gene flow between *S. latifolia* and *S. dioica* (Minder et al. 2007; Karrenberg and Favre 2008), this might affect different genome regions differently (Muir et al. 2012). Before comparing associations between the partially

sex-linked genes and the MSY in the two species, we therefore tested whether the PAR gene sequences we obtained might show less gene flow than autosomal genes. Based on all site types, divergence is invariably low, and the PAR genes studied here have mean net divergence, D_a , of 0.25%, and include no fixed differences between the species, while shared variants are common, similar to the published results for autosomal and X-linked genes (Supplementary Table S2), and consistent with incomplete isolation between the two species. Net divergence, D_a , ranges from only 0.55% for 20 autosomal genes (Muir et al. 2012), or 0.27% in a larger sample of genes recently studied in three populations of each species (Hu and Filatov 2015), to 0.51% for X-linked genes (Hu and Filatov 2015). K_{ST} values for the 6 PAR genes studied here nevertheless indicate differentiation between *S. latifolia* and *S. dioica* (Supplementary Table S2).

Gene flow between *S. latifolia* and *S. dioica* does not appear to be highly restricted for the PAR genes studied here. One PAR gene, *E521*, had high diversity in *S. latifolia*, because many sites are heterozygous in all males, reflecting diverged sequences associated with the X and Y chromosome fully sex linked haplotypes (Qiu et al. 2016), but the X-linked alleles showed no variants. In *S. dioica*, we detected no variation in *E521*, despite testing several pairs of PCR primers to exclude the possibility that a highly diverged allele is present and is not amplified. Excluding *E521*, whose low diversity in *S. dioica* will increase K_{ST} , the mean K_{ST} value for the five other PAR genes studied here is 0.084 (with 95% confidence intervals 0.034-0.128), versus the estimated values of 0.154 for autosomal genes, or 0.381 for X-linked genes (Hu and Filatov 2015).

Subdivision of PAR boundary genes between males and females in *S. latifolia* versus *S. dioica*

Our results suggest that, consistent with the cytogenetic evidence outlined above, the genes in the added regions closely linked to the *S. latifolia* PAR-MSY boundary are also partially sex-linked in *S. dioica*. However, they appear to recombine more often in *S. dioica*. A first indication of sex linkage is higher nucleotide diversity in males than in females, indicating that some variants in the

sequences are male-specific. The difference is significant in *S. latifolia* (Mann-Whitney U-test, $P = 0.03$), and many variants are present only in males, or with much higher frequencies in males than females (Supplementary Figures S2 and S3). A significantly higher frequency of sites that are heterozygous in males than in females (Mann-Whitney U-test, $P = 0.012$), with many sites heterozygous in most males, but few in females (Supplementary Figures S2-S4), shows that the sex difference in diversity is not due to our sample having included some males from a population with divergent sequences, consistent with recombinant alleles not being geographically restricted (Qiu et al. 2016). In *S. dioica*, however, diversity is only slightly higher in males (Figure 2), though male-specific variants in gene *E559* (Supplementary Figure S2), and high heterozygote frequencies in males in several of the genes studied (Figure 3A, Supplementary Figure S5) suggest partial sex linkage of at least some of these genes in this species also. However, many variants are not strongly associated with the fully sex-linked region (Supplementary Figures S2 and S3); again recombinant alleles are not confined to any geographic region or sub-set of populations.

Figure 2 and Figure 3 about here

Tajima's D values more clearly suggest sex linkage of some of these genes in both species (Figure 3B). In *S. latifolia*, D values are positive for most PAR genes in males, whereas other genome regions consistently show negative values (Qiu et al. 2016). For the genes studied here, males and females differ significantly ($D_{\text{males}} = 1.28$, $D_{\text{females}} = -1.29$, Mann-Whitney U-test, $P = 0.02$), suggesting that these genes are closely linked with the fully sex-linked region, so that variants are often heterozygous in males. Positive Tajima's D values in males, and negative values in females are also seen in *S. dioica*, particularly for genes *cs3597* and *cs935*, where the value is statistically significant in males (Figure 3B and Table S3).

As explained above, analysis of subdivision between the sexes is the best way to test for associations of variants with the fully sex-linked region (Kirkpatrick and Guerrero 2014). For all five PAR boundary genes that could be tested in *S. latifolia*, K_{ST} between sequences from males and females was high, and differed highly

significantly from zero (Table 1), unlike the more distally located *cs4991* gene ($P=0.052$), or other PAR genes more loosely linked to the fully sex-linked region, or autosomal loci (Qiu et al. 2010). Consistent with partial sex linkage, the significant associations in *S. latifolia* are somewhat weaker than the value of one third expected for fully sex-linked genes assuming a 1:1 sex ratio and fixed differences between Y- and X linked alleles. In *S. dioica*, the K_{ST} values were smaller, but significant subdivision between the sexes was detected for several PAR boundary genes (Table 1). Because variances are not available for K_{ST} values, we tested for subdivision by two significance tests, and by comparing Tajima's D values between the sexes, including using sub-samples of sequences from each species; for three genes, at least one test suggests partial sex linkage in *S. dioica* (Supplementary Figure S4).

Table 1 about here

The lower significance levels for our tests of subdivision between the sexes in *S. dioica* are not wholly due to our having sequenced fewer alleles than from *S. latifolia*, because smaller sub-samples of sequences from the *S. latifolia* alleles, re-analyzed in the same manner as for the complete data set (see Methods) consistently yielded K_{ST} values much higher than in *S. dioica*, and highly significant subdivision between the sexes (Table 1). For two of the genes, *cs3597* and *E523*, none of our 40 sub-samples had K_{ST} value as low as the *S. dioica* value. For the other genes, only a few sub-samples from *S. latifolia* had lower K_{ST} values than those observed in *S. dioica*, and the means over the 40 sub-samples are invariably higher (Table 1).

Sub-samples also maintained the higher Tajima's D values in males than females in *S. latifolia* (Supplementary Table S3). Although the difference between the sexes is smaller in the sub-samples than in the full data set, as expected, it remains statistically significant (Mann-Whitney U-test, $P = 0.03$). Overall, the results from the sub-samples show that there is a real biological difference between the species.

Although only a 120 bp region of *E521* could be sequenced in *S. latifolia* females, and in *S. dioica* (Supplementary Table S2), our results also support the conclusion that this gene is also closely associated with the MSY in *S. latifolia* (Figures 1 to 3 and Supplementary Table S3), but not in *S. dioica*, in which there is a complete absence of variants, and thus no difference between the sexes.

As explained above, the *cs3597* gene is probably part of an ancestral PAR (Qiu et al. 2016), and appears partially sex-linked in both species based on Tajima's *D* values (Figure 3B) and *K_{ST}* analysis. Interestingly, even though this gene was probably partially sex-linked in the common ancestor, it has a considerably smaller allele frequency difference and *K_{ST}* between the sexes in *S. dioica* than *S. latifolia* (see Table 1), and a smaller difference in the frequency of heterozygotes (Figure 3A). It has therefore probably become more closely linked to the MSY in *S. latifolia*.

Evidence that closer linkage has evolved in *S. latifolia* since splitting from *S. dioica*

To infer more rigorously whether the recombination state is changed in *S. latifolia*, or in *S. dioica*, an outgroup is needed. We therefore sequenced two genes in *S. marizii* and *S. diclinis*, one from each of the two additions that formed the PAR (see above). Both these genes include many sites showing complete sex linkage in *S. latifolia*, but neither of them included any fully Y-linked variants in *S. marizii* or *S. diclinis* (Supplementary Figures S2 and S3). These genes are therefore not completely sex linked in *S. marizii* and *S. diclinis*; they could be partially sex linked, but, because we have only small samples of these two species (see Supplementary Table S1), we have no firm evidence that the genes studied are not autosomal in these species. Both the rearrangements that created the PAR in *S. latifolia*, and the one that is firmly inferred in *S. dioica*, might thus have occurred after the split from *S. marizii* and *S. diclinis*. Nevertheless, taken together, these results suggest that the *S. latifolia* state of strong associations with the fully sex-linked region is the derived state for PAR boundary genes derived from both addition events that formed the PAR, and that the other three species share the less closely linked state.

DISCUSSION

Overall, the difference between *S. latifolia* versus *S. dioica* in associations of sequence variants with the sexes (reflecting linkage disequilibrium between alleles of PAR boundary region genes and the MSY region), suggests less recombination between these PAR boundary genes in *S. latifolia* than in *S. dioica*. Subdivision between X and Y chromosomes due to suppressed recombination between the XY pair in males will be most evident when X and Y haplotypes are sequenced, and less readily detected in samples of sequences from males and females (Kirkpatrick and Guerrero 2014), so that some discrepancies between different tests for associations between the MSY and gene sequences are unsurprising. However, for partially sex-linked genes, the phase of variants is not known. Our approach of testing sequences from males and females, without attempting to infer their phases, is conservative.

Our analysis suggests recently decreased recombination in *S. latifolia*, but this could be either by suppression of recombination between the MSY and the region carrying the genes we studied, or because the translocations that brought these genes onto the XY pair occurred in the short evolutionary time since *S. latifolia* split from *S. dioica*. We consider this less likely, as our analyses of subdivision between males and females suggests partial sex-linkage in *S. dioica*, for at least some of the PAR boundary genes that we tested, implying addition before the split between the species. This is consistent with cytogenetic evidence that both the X and Y chromosomes are very similar in these species (Grabowska-Joachimciak and Joachimciak 2002), and considerably larger than other chromosomes, unlike the situation in other *Silene* species (Siroky et al. 2001).

Two of the genes studied, *E523* and *cs935*, may not be sex-linked in *S. dioica* (see Table 1), though subdivision between the two sexes is supported for *cs935* by a significant result ($P = 0.2\%$) with the *Snn* test of Hudson et al. (1992), and a large positive Tajima's D value in males but not females, see Figure 3, suggesting partial sex linkage; both these genes map to the SvLG9 linkage group in *S. vulgaris*, and were probably added to the evolving *S. latifolia* PAR in the second of the two inferred addition events diagrammed in Figure 1 (Qiu et al. 2016). This event could potentially have occurred after the split between the two species, leaving these genes autosomal in *S. dioica* (this could be tested by in situ hybridisation of genes from this linkage group). Alternatively, this event pre-dated the split, but close

linkage has subsequently evolved very recent, and only in *S. latifolia*. Interestingly, however, genes *E521* and *E559* are located on LGSmall of *S. vulgaris*, indicating that they became part of a PAR in the first translocation event (Qiu et al. 2016), yet only *E559* shows associations with the MSY in *S. dioica*.

Linkage disequilibrium, and associations reflecting it, depends on the effective population size (N_e) as well as the recombination rate, and also on natural selection. However, the difference we find between *S. latifolia* and *S. dioica* cannot be explained by an N_e difference. The estimated silent site nucleotide diversity values for autosomal genes differ very slightly between *S. latifolia* and *S. dioica*, based on estimates from large numbers of genes (Hu and Filatov 2015). Diversity estimates for silent sites are most appropriate for assessing effective population sizes, and the largest difference in such diversity estimates so far published is 7% (in the estimated value of π in Muir et al. 2012), which is too small to explain the observed difference in associations of variants in the PAR boundary genes with the MSY between *S. latifolia* and *S. dioica* (the alternative diversity estimate in the same paper, Watterson's theta, yielded a slight difference in the opposite direction). The difference in K_{ST} between males and females in the two species therefore probably reflects different recombination rates between the MSY and the genes studied.

Because only one of the genes studied is part of the ancestral PAR, while the other four became sex-linked through the translocation events, it is likely that linkage has become closer in *S. latifolia* and that *S. dioica* remained more similar to the ancestral state. This is supported by the lack of evidence for sex linkage in the two outgroup species, *S. marizii* and *S. diclinis*. Whether the translocations are shared between all four species studied here, or whether one addition or both is present in *S. dioica*, closer linkage has nevertheless clearly evolved subsequently, at least in *S. latifolia*.

Situations where regions have been added to pre-existing sex chromosomes are of great interest, because one factor that has been proposed to favour such rearrangements is sexual antagonism (Charlesworth and Charlesworth 1980; Pennell et al. 2015), similar to the selection leading to new evolutionary strata on sex chromosomes in the first instance (see Introduction). This view predicts that closer

linkage with the sex-determining region should subsequently evolve, unless the rearrangement directly caused close linkage between the sex-determining region and the SA factor. If closer linkage with the sex-determining region is generally found to evolve, this would lend support to the SA polymorphism hypothesis. We therefore next briefly review the evidence about whether such changes are, in fact, seen when genome regions have been added to sex chromosomes of species other than *Silene*.

Recombination in genome regions added to regions to sex chromosomes

In XY species with no recombination in males, such as *Drosophila*, Y-autosome and X-autosome translocations both immediately result in complete sex-linkage (reviewed in Bachtrog 2006). While these have been important for studying the consequences of suppressed recombination (reviewed in Bachtrog 2006), they are uninformative regarding the evolution of recombination and the possible involvement of sexual antagonism.

In species with recombination in the heterogametic sex, however, chromosomes can initially continue recombining in both sexes, even after Y-autosome translocations in XY species, or W-autosome translocations in ZW species like birds, where females are the heterogametic sex; similarly, X-A translocations in systems with XO males form neo-Y chromosomes that can continue to recombine with their autosomal counterparts (e.g. Castillo et al. 2010; Henzel et al. 2011). It is therefore interesting to ask whether recombination suppression evolves subsequently.

Additions to fully sex-linked regions

In *Silene diclinis* the neo-Y (Y₂) arm created by a Y-autosome reciprocal translocation probably largely recombines with the former autosome (Howell et al. 2009), although population genetic studies to confirm this have not yet been done, nor genetic mapping to study recombination patterns in detail. Similarly, some threespine stickleback populations have a Y-A translocation involving the fully sex-linked end of the Y chromosome (Natri et al. 2013). In this case, however, a region adjacent to the rearrangement breakpoint shows large allele frequency differences

between the sexes, indicating suppressed recombination (reviewed in Natri et al. 2013). Sex linkage of this region, extending for least 3 Mb, could simply be due to the rearrangement, as a neo-X or –Y chromosome that segregates from an enlarged Y or X, but is not physically joined to it, may fail to pair near the breakpoints.

In yet other cases, however, recombination has probably subsequently become suppressed. For example, the neo-Y chromosomes have become heterochromatic in several groups of related species with shared Robertsonian fusions. Examples include grasshoppers (Castillo et al. 2010) and deer (Cernohorska et al. 2015).

Sex chromosome-autosome translocations have occurred in several dioecious plants (reviewed in Ming et al. 2011), and these may be excellent for studying whether recombination suppression has evolved, because the pre-rearrangement state is often known in outgroup species, as in *S. diclinis*. In *Rumex acetosa*, the Y₂ is heterochromatic, suggesting suppressed recombination (however, the involvement of an X-autosome translocation is not currently certain in this plant, see Rejón et al. 1994). Several other plants with XX/XY₁Y₂ systems probably have new regions added to fully X-linked regions through X-autosome translocations, so that recombination could be suppressed in males in parts of the added regions as direct effects of the rearrangements. In the hop species *Humulus lupulus* var. *cordifolius* and *H. japonicus*, the Y₂ (neo-Y, former autosome) chromosome is still non-heterochromatic (Grabowska-Joachimik et al. 2011). The situation is similar in the genus *Baccharis* (Hunziker et al. 2002). In *Viscum fischeri*, a chain of five Y chromosomes (Barlow and Wiens 1976) suggests multiple events over considerable evolutionary time, so that multiple strata, corresponding to different translocation events might exist. X-Y divergence data could indicate whether the neo-Y chromosomes of these plants still recombine.

Additions to partially sex-linked regions

Few cases of additions onto PARs have been studied, although possible cases have been inferred in birds. In two warblers, markers from a previously autosomal micro-chromosome (chicken 4a) have alleles on both the Z and W chromosomes

(Pala et al. 2012), and larks may be similar (Brooke et al. 2010). However, it has not yet been excluded that these additions occurred onto the fully sex-linked region. This could be tested by genetic maps of these species and outgroup species lacking the rearrangement, and tests for complete sex-linkage of the added genes, and other genes from the chicken linkage groups involved in these events, would be informative.

Cases with strong evidence for translocations that added new regions onto a PAR, as in *S. latifolia* and its close relatives, are thus particularly interesting. Other than *S. latifolia*, an addition to the PAR has so far been suggested in only one other plant, *Rumex hastatulus* (Grabowska-Joachimciak et al. 2015). This species is in the same dioecious clade as *R. acetosa* (Navajas-Pérez et al. 2005), but the rearrangements may be independent, as the *R. hastatulus* one is found only in one race. This North Carolina race has an XX/XY₁Y₂ sex chromosome system, so an X-autosome fusion (Smith 1964) is more likely than an addition to the PAR. The neo-Y has remained euchromatic (Grabowska-Joachimciak et al. 2015), consistent with a smaller divergence of its sequences from those of the neo-X, and less signs of genetic degeneration, compared with the ancestral chromosome pair found in the closely related XX/XY Texas race (Hough et al. 2014). Moreover, while independently sequenced transcriptomes from six populations of each race found male-specific variants in ~80% of the genes ascertained as sex-linked in the XY populations, this was found for only 28% of the neo-Y genes, suggesting that most of the Y₂ still recombines (Hough et al. 2014). As for the species discussed in the previous section, it remains unclear whether any of the ancestrally autosomal genes have become fully sex-linked.

What might cause a difference in recombination between *S. latifolia* and *S. dioica*?

The mechanism causing the apparent difference in recombination between *S. dioica* and *S. latifolia* is currently unknown, and it is not yet clear whether a single change was involved, such as lower crossing over in a single region proximal to the fully sex-linked region (leaving the genetic map distances of most intervals unchanged), or several changes. In *S. latifolia*, the PAR has undergone

rearrangements, but the current order of the PAR boundary region genes appears to be the same in the X and Y (Qiu et al. 2016), which suggests that they became closely linked to the *S. latifolia* MSY after the rearrangements, which therefore probably did not cause the changed recombination.

We can also probably exclude the inversion that has been detected in the *S. dioica* X chromosome. This inversion is not shared with the putative outgroup species, *S. diclinis*, so *S. dioica* is thought to have the derived state (Nicolas et al. 2005); moreover, the inversion is probably confined to *S. dioica*, because the order of the genes that allowed the inversion to be detected differs from that in *S. latifolia*, which is the same as that in the more distant outgroup *S. vulgaris*. The extent of the inversion is not yet known, as these genes belong to the younger stratum (stratum II of Bergero et al. 2013), and it could extend some way into the older stratum, and/or into the PAR. The difference we detect affects recombination between PAR and fully Y-linked variants, and must therefore be due to a difference in male meiosis. If the inversion lies wholly within the X-linked region, recombination in regions near the PAR boundary would probably not be affected in males. If, however, it extends into the PAR, recombination events in *S. dioica* heterozygotes might shift to more distal parts of the chromosome (Henzel et al. 2011), causing genes that are closely linked to the *S. latifolia* PAR boundary to show complete sex-linkage in *S. dioica*, the opposite of the effect we observe. Finally, our limited results from *S. diclinis* (probably lacking the inversion) show that the orthologues of the genes studied here probably recombine with the MSY, as in *S. dioica*, contrasting with the closer linkage in *S. latifolia*.

It also seems unlikely that a chromosome rearrangement such as an inversion could be involved, because the PAR gene sequences appear to recombine in both *S. dioica* and *S. latifolia*. All the PAR boundary genes we sequenced included variants shared between *S. latifolia* and *S. dioica* (consistent with their low K_{ST} values between these species, see above). This indicates that gene flow between the species occurs in the genome region that we studied, arguing against an inversion, since a rearrangement would impede gene flow. Moreover, the *S. dioica* sequences containing variants shared with *S. latifolia* are not intact, unrecombined copies of *S.*

latifolia haplotypes (see Figures S1 and S2). Rather, a few sites in the *S. dioica* sequences show clear signs of sex-linkage, with variants appearing either X- or Y-linked, while other sites in the same genes have the variant that appears X-linked in *S. latifolia*, or the Y-linked variant.

Overall, therefore, we conclude that the PAR boundary genes studied here probably did not immediately become fully sex-linked when they were added to the sex chromosomes, but that reduced recombination has subsequently evolved, so that the region now recombines rarely with the Y-linked region, particularly in *S. latifolia*.

What might have led to the difference in recombination between the closely related species *S. latifolia* and *S. dioica*? It is possible that some general pressure promotes the evolution of reduced recombination. As outlined above, one possibility is that the ultimate cause is a selective pressure due to a sexually antagonistic polymorphism at a locus in the genome region close to the PAR boundary (requiring that the gene and the loci studied here are very closely linked to the boundary, as reviewed in Qiu et al. 2016); the observed allele frequency differences between the sexes are consistent with the presence of such a polymorphism in *S. latifolia*. This variant might not have become established in *S. dioica*, explaining the difference. However, an equally plausible alternative is that such a polymorphism arose in an ancestor of both species, but a response to the selection for closer linkage occurred in only one species, due to a lack of genetic variation for the recombination rate in this region in the other. Alternatively, the change in recombination could be caused by a non-selective force. For example, it has been hypothesized that fully sex-linked regions may expand, at the expense of PARs, by an automatic process that redistributes crossovers towards the distal regions of sex bivalents as the sex-specific region differentiates and heteromorphism evolves, causing chromosome asymmetry (Henzel et al. 2011). Again, this might differ between closely related species.

Understanding the proximate mechanism(s) involved may help to distinguish between these different possibilities leading to suppressed recombination between sex chromosomes. The presence of sexually antagonistic polymorphisms may favour chromosome rearrangements or other major factors controlling recombination and preventing recombination across regions carrying many genes (Charlesworth and

Charlesworth 1980), while the alternative just outlined, which could occur without SA effects being involved, might cause gradual extension of the border of the non-recombining region into the PAR. Our results establish *S. latifolia* as a species in which it may be possible to learn in detail about changes that are extending the sex chromosomes' non-recombining regions, at the expense of the PAR.

ACKNOWLEDGEMENTS: We thank the Leverhulme Trust for funding (RPG-2013-110). S. G-R was supported by a Beatriu de Pinós postdoctoral fellowship (AGAUR; 2010 BP_A 00438). We thank Sophia Ahmed, Sophie Karrenberg and Kai Zeng for samples of *S. dioica*, Lynda Delph for samples of *S. diclinis*, and Cristina Vieira and Francisco Amich for samples of *S. marizii*. We are grateful to the Edinburgh University greenhouse staff for excellent plant care.

DATA ARCHIVING: The sequence data are available in Dryad under accession number [TO BE ADDED].

Supplementary information is available at (the journal's name TO BE ADDED)'s website

CONFLICT OF INTEREST: The authors declare no conflict of interest.

LITERATURE CITED

- Bachtrog D (2006) A dynamic view of sex chromosome evolution. *Curr. Opin. Genet. Dev.* **16**: 578-585.
- Barlow BA, and Wiens D (1976) Translocation heterozygosity and sex ratio in *Viscum fischeri*. *Heredity* **37**: 27-40.
- Bergero R, Forrest A, Kamau E and Charlesworth D (2007) Evolutionary strata on the X chromosomes of the dioecious plant *Silene latifolia*: evidence from new sex-linked genes. *Genetics* **175**: 1945-1954.
- Bergero R, Qiu S, Forrest A, Borthwick H and Charlesworth D (2013) Expansion of the pseudoautosomal region and ongoing recombination suppression in the *Silene latifolia* sex chromosomes. *Genetics* **194**: 673-686.

- 650 Brooke M, Welbergen J, Mainwaring M, Van-Der-Velde M, Harts A *et al.* (2010)
 651 Widespread translocation from autosomes to sex chromosomes preserves
 652 genetic variability in an endangered lark. *J. Mol. Evol.* **70**: 242-246.
- 653 Bull JJ, (1983) *Evolution of Sex Determining Mechanisms*. Benjamin/Cummings,
 654 Menlo Park, CA.
- 655 Castillo E, Bidau C and Marti D (2010) Neo-sex chromosome diversity in Neotropical
 656 melanopline grasshoppers (Melanoplineae, Acrididae). *Genetica* **138**: 775-786.
- 657 Cernohorska H, Kubickova S, Kopečna O, Vozdova M, Matthee C *et al.* (2015) Nanger,
 658 Eudorcas, Gazella, and Antelope form a well-supported chromosomal clade
 659 within Antilopini (Bovidae, Cetartiodactyla). *Chromosoma* **124**: 235-247.
- 660 Charlesworth B, Nordborg M and Charlesworth D (1997) The effects of local
 661 selection, balanced polymorphism and background selection on equilibrium
 662 patterns of genetic diversity in subdivided inbreeding and outcrossing
 663 populations. *Genet. Res.* **70**: 155-174.
- 664 Charlesworth D, and Charlesworth B (1980) Sex differences in fitness and selection
 665 for centric fusions between sex-chromosomes and autosomes. *Genetical*
 666 *Research (Cambridge)* **35**: 205-214.
- 667 Chibalina M, and Filatov D (2011) Plant Y chromosome degeneration is retarded by
 668 haploid purifying selection. *Curr. Biol.* **21**: 1475–1479.
- 669 Cortez D, Marin R, Toledo-Flores D, Froidevaux L, Liechti A *et al.* (2014) Origins and
 670 functional evolution of Y chromosomes across mammals. *Nature* **508**: 488–493.
- 671 Cotter DJ, Brotman SM and Sayres MW (2016) Genetic diversity on the human X
 672 chromosome does not support a strict pseudoautosomal boundary. *Genetics*
 673 **203**: 485-492.
- 674 Grabowska-Joachimik A, and Joachimik A (2002) C-banded karyotypes of two
 675 *Silene* species with heteromorphic sex chromosomes. *Genome* **45**: 243-252.
- 676 Grabowska-Joachimik A, Kula A, Książczyk T, Chojnicka J, Sliwiska E *et al.* (2015)
 677 Chromosome landmarks and autosome-sex chromosome translocations in
 678 *Rumex hastatulus*, a plant with XX/XY1Y2 sex chromosome system. *Chromosome*
 679 *Res.* **23**: 187-197.
- 680 Grabowska-Joachimik A, Mosiolek M, Lech A and Góralski G (2011) C-Banding/DAPI
 681 and *in situ* hybridization reflect karyotype structure and sex chromosome

- 682 differentiation in *Humulus japonicus* Siebold & Zucc. *Cytogenet Genome*
 683 *Research* **132**: 203–211.
- 684 Henzel JV, Nabeshima K, Schvarzstein M, Turner B, Villeneuve AM *et al.* (2011) An
 685 asymmetric chromosome pair undergoes synaptic adjustment and crossover
 686 redistribution during *Caenorhabditis elegans* meiosis: implications for sex
 687 chromosome evolution. *Genetics* **197**: 685–699.
- 688 Hough J, Hollister JD, Wang W, Barrett SCH and Otto SP (2014) Genetic degeneration
 689 of old and young Y chromosomes in the flowering plant *Rumex hastatulus*. *Proc.*
 690 *Natl. Acad. Sci. U. S. A.* **111**: 7713–7718.
- 691 Howell EC, Armstrong S and Filatov D (2009) Evolution of neo-sex chromosomes in
 692 *Silene diclinis* *Genetics* **182**: 1109 - 1115.
- 693 Hu X, and Filatov DA (2015) The large-X effect in plants: Increased species divergence
 694 and reduced gene flow on the *Silene* X-chromosome. *Mol. Ecol.*
- 695 Hudson RR, Boos DD and Kaplan NL (1992) A statistical test for detecting geographic
 696 subdivision. *Mol. Bio. Evol.* **9**: 138-151.
- 697 Hunziker J, Wulff A and Escobar A (2002) Permanent translocation heterozygosity in
 698 dioecious *Baccharis coridifolia* DC. (Asteraceae). *Hereditas* **137**: 132-139.
- 699 Jordan C, and Charlesworth D (2012) The potential for sexually antagonistic
 700 polymorphism in different genome regions. *Evolution* **66**: 505–516.
- 701 Kaiser VB, Bergero R and Charlesworth D (2009) *SlCyt*, a newly identified sex-linked
 702 gene, has recently moved onto the X chromosome in *Silene latifolia*
 703 (Caryophyllaceae). *Mol. Bio. Evol.* **26**: 2343 - 2351.
- 704 Karrenberg S, and Favre A (2008) Genetic and ecological differentiation in the
 705 hybridizing champions *Silene dioica* and *S. latifolia*. *Evolution* **62**: 763-773.
- 706 Kirkpatrick M, and Guerrero R (2014) Signatures of sex-antagonistic selection on
 707 recombining sex chromosomes. *Genetics* **197**: 531-541.
- 708 Laporte V, Filatov DA, Kamau E and Charlesworth D (2005) Indirect evidence from
 709 DNA sequence diversity for genetic degeneration of Y-chromosome in dioecious
 710 species of the plant *Silene*: the *SIY4/SIX4* and *DD44-X/DD44-Y* gene pairs. *J. Evol.*
 711 *Biol.* **18**: 337-347.
- 712 Librado P, and Rozas J (2009) DnaSP v5: a software for comprehensive analysis of
 713 DNA polymorphism data. *Bioinformatics* **25**: 1451-1452.

- 714 Minder A, Rothenbuehler C and Widmer A (2007) Genetic structure of hybrid zones
 715 between *Silene latifolia* and *Silene dioica* (Caryophyllaceae): evidence for
 716 introgressive hybridization. *Mol. Ecol.* **16**: 2504-2516.
- 717 Ming R, Bendahmane A and Renner S (2011) Sex chromosomes in land plants. *Annual*
 718 *Review of Plant Biology* **62**: 485-514.
- 719 Muir G, Dixon CJ, Harper AL and Filatov DA (2012) Dynamics of drift, gene flow and
 720 selection during speciation in *Silene*. *Evolution* **66**: 1447–1458.
- 721 Natri HM, Shikano T and Merilä J (2013) Progressive recombination suppression and
 722 differentiation in recently evolved neo-sex chromosomes. *Mol. Bio. Evol.* **30**:
 723 1131-1144.
- 724 Navajas-Pérez R, Herran RDL, Gonzalez G, Jamilena M, Lozano R *et al.* (2005) The
 725 evolution of reproductive systems and sex-determining mechanisms within
 726 *Rumex* (Polygonaceae) inferred from nuclear and chloroplastidial sequence data.
 727 *Mol. Bio. Evol.* **22**: 1929-1939.
- 728 Nicolas M, Marais G, Hykelova V, Janousek B, Laporte V *et al.* (2005) A gradual
 729 process of recombination restriction in the evolutionary history of the sex
 730 chromosomes in dioecious plants. *PLoS Biol.* **3**: 47-56.
- 731 Pala I, Naurin S, Stervander M, Hasselquist D, Bensch S *et al.* (2012) Evidence of a
 732 neo-sex chromosome in birds. *Heredity* **108**: 264–272.
- 733 Papadopoulos AST, Chester M, Ridout K and Filatov DA (2015) Rapid Y degeneration
 734 and dosage compensation in plant sex chromosomes. *Proc. Natl. Acad. Sci. U. S.*
 735 *A.* **112**: 13021–13026.
- 736 Pennell M, Kirkpatrick M, Otto S, Vamosi J, Peichel C *et al.* (2015) Y fuse? Sex
 737 chromosome fusions in fishes and reptiles. *PLoS Genetics* **11**: e1005237.
- 738 Prentice HC (1978) Experimental taxonomy of *Silene* section *Elisanthe*
 739 (Caryophyllaceae): crossing experiments. *Bot. J. Linn. Soc.* **77**: 203–216.
- 740 Qiu S, Bergero R and Charlesworth D (2013) Testing for the footprint of sexually
 741 antagonistic polymorphisms in the pseudo-autosomal region of a plant sex
 742 chromosome pair. *Genetics* **194**: 663-672.
- 743 Qiu S, Bergero R, Forrest A, Kaiser VB and Charlesworth D (2010) Nucleotide diversity
 744 in *Silene latifolia* autosomal and sex-linked genes. *Proc. R. Soc. B* **277**: 3283-3290.

- 745 Qiu S, Bergero R, Guirao-Rico S, Campos JL, T Cezard *et al.* (2016) RAD-mapping
 746 reveals an evolving, polymorphic and fuzzy boundary of a plant
 747 pseudoautosomal region. *Mol. Ecol.* **25**: 414–430
- 748 Rautenberg A, Hathaway L, Oxelman B and Prentice HC (2010) Geographic and
 749 phylogenetic patterns in *Silene* section *Melandrium* (Caryophyllaceae) as
 750 inferred from chloroplast and nuclear DNA sequences. *Mol. Phylogenet. Evol.* **57**:
 751 978-991.
- 752 Rejón CR, Jamilena M, Ramos MG, Parker JS and Rejón MR (1994) Cytogenetic and
 753 molecular analysis of the multiple sex-chromosome system of *Rumex acetosa*.
 754 *Heredity* **72**: 209-215.
- 755 Rice WR (1987) The accumulation of sexually antagonistic genes as a selective agent
 756 promoting the evolution of reduced recombination between primitive sex-
 757 chromosomes. *Evolution* **41**: 911-914.
- 758 Siroky J, Lysak MA, Dolezel J, Kejnovsky E and B.Vyskot (2001) Heterogeneity of rDNA
 759 distribution and genome size in *Silene* spp. *Chromosome Res.* **9**: 387-393.
- 760 Smith BW (1964) The evolving karyotype of *Rumex hastatulus*. *Evolution* **18**: 93-104.
- 761 Städler T, Haubold B, Merino C, Stephan W and Pfaffelhuber P (2009) The impact of
 762 sampling schemes on the site frequency spectrum in nonequilibrium subdivided
 763 populations. *Genetics* **182**: 205-216.
- 764 Tajima F (1989) Statistical method for testing the neutral mutation hypothesis.
 765 *Genetics* **123**: 585-595.
- 766 Wang J, Na J, Yu Q, Gschwend AR, Han J *et al.* (2012) Sequencing papaya X and Yh
 767 chromosomes reveals molecular basis of incipient sex chromosome evolution.
 768 *Proc. Natl. Acad. Sci. U. S. A.* **109**: 13710-13715.
- 769 Zhou Q, Zhang J, Bachtrog D, An N, Huang Q *et al.* (2014) Complex evolutionary
 770 trajectories of sex chromosomes across bird taxa. *Science* **346**: 1332.

Titles and legends to figures

Figure 1. Evolution of *S. latifolia* PAR, showing the origins of the PAR boundary genes studied in this paper (simplified from Figure 4 of Qiu et al. 2016, and including gene locations only for these PAR boundary genes). **A.** The ancestral state before the translocation events had a sex chromosome with a PAR (green region) and a region that had stopped recombining (indicated in black region, but not as separate X-linked and male-specific, or MSY, regions; different strata are not indicated in this diagram, but the black line includes all fully sex-linked regions). The two autosomes that later became translocated, moving the whole or part of these chromosome arms onto the PAR of the ancestral sex chromosome, are represented as pink and blue lines. The first and second translocation events are represented in parts **B** and **C**, with pink and blue indicating the linkage groups in *S. vulgaris* that carry the PAR boundary genes, when this is known. The *cs4991* gene is not indicated because it could not be mapped in our *S. vulgaris* family; however its location in the genetic map of the *S. latifolia* X chromosome is among the more distal genes added in the second translocation event. Other genes added in both translocations have remained loosely linked to the MSY boundary, and are not shown in the diagram.

Figure 2. Comparisons of nucleotide diversity in males and females of *S. latifolia* (black and grey) and *S. dioica* (dark and pale pink). The estimated diversity values for silent sites are shown above the x axis, and the differences in diversity between males and females are shown below the axis. Note that there are no values for gene *E521* in *S. dioica* because no variants were found in the sequenced region of this gene in this species.

Figure 3. Heterozygote frequencies at polymorphic sites (A) and Tajima's *D* values (B) for all site types in males and females of *S. latifolia* (black and grey, respectively) and *S. dioica* (dark and pale pink). In panel A, the differences in heterozygote frequencies between males and females are also shown.

803 **Supplementary files**

804 **Table S1.** Locations of the populations from which the *S. dioica*, *S. diclinis* and *S.*
805 *marizii* plants were collected.

806 **Table S2.** Divergence and K_{ST} estimates for *S. dioica* and *S. latifolia*.

807 **Table S3.** Tajima's D values in *S. dioica* and *S. latifolia*. For *S. latifolia*, results are
808 shown for the complete set of sequences, and for sub-sampled sets (see text).

809 **Figure S1.** Locations from which the plant samples were collected.

810 **Figure S2.** Non-singleton variants in gene *E559* in the four *Silene* species studied.

811 **Figure S3.** Non-singleton variants in gene *E523* in the four *Silene* species studied.

812 **Figure S4.** K_{ST} values with error bars showing the ranges of values in 40 sub-samples
813 of sequences, as described in the Methods section of the main text.

814 **Figure S5.** Heterozygote frequencies at non-singleton sites in 5 genes in *S. latifolia*
815 and *S. dioica* males (blue) and females (pink). Gene *E521* is not shown, because there
816 were no polymorphic sites in *S. dioica*. The other genes are shown in the order of
817 their genetic map locations on the X chromosome, as inferred by Qiu et al. (2015),
818 with *cs3597* closest to the fully sex-linked region.

Table 1. K_{ST} values and tests of subdivision between males and females in complete sequence sets and in sub-samples from *S. latifolia* with the same size as the *S. dioica* data. For the *S. dioica* results, P values are given for the significance tests using K_{ST}^* with 1000 permutations (Hudson et al. 1992), and significance is indicated by bold text; in *S. latifolia*, all the PAR boundary genes, but not the more distal *cs4991* gene, have K_{ST} values significantly different from zero with $P < 0.0001$. *S. latifolia* sub-samples were not created for gene *E521*, where there are no variants in *S. dioica*, so that we cannot compare the significance of subdivision between males and females, or for *cs4991*, where neither species has a K_{ST} value significantly different from zero.

Gene	<i>S. vulgaris</i> linkage group ¹	Sequence length ²	K_{ST} values in complete sequence data sets, and significance test results			Sub-samples from <i>S. latifolia</i> data		
			<i>S.</i> <i>latifolia</i> K_{ST}	<i>S. dioica</i>		Mean K_{ST} in 40 sub- samples	Numbers of sub- samples with $K_{ST} \leq$ <i>S. dioica</i> value	Number of significant values in sub-samples (number with $P < 0.01$)
				K_{ST}	P values			
<i>cs3597</i>	SvLG12	177	0.365	0.08	0.023	0.292	0	40 (40)
<i>E559</i>	SvLGSmall	382	0.194	0.079	0.003	0.168	1	40 (24)
<i>E521</i>	SvLGSmall	120	0.213	— (no variants)		—	—	—
<i>E523</i>	SvLG9	486	0.202	0.032	0.057	0.206	0	40 (40)
<i>cs935</i>	SvLG9	388	0.071	0.02	0.13	0.064	1	28 (12)
<i>cs4991</i>	Not known	282	0.0158	0.00262	0.354	—	—	—

¹ From Qiu et al. 2015

² Number of sites excluding alignment gaps

Figure 1

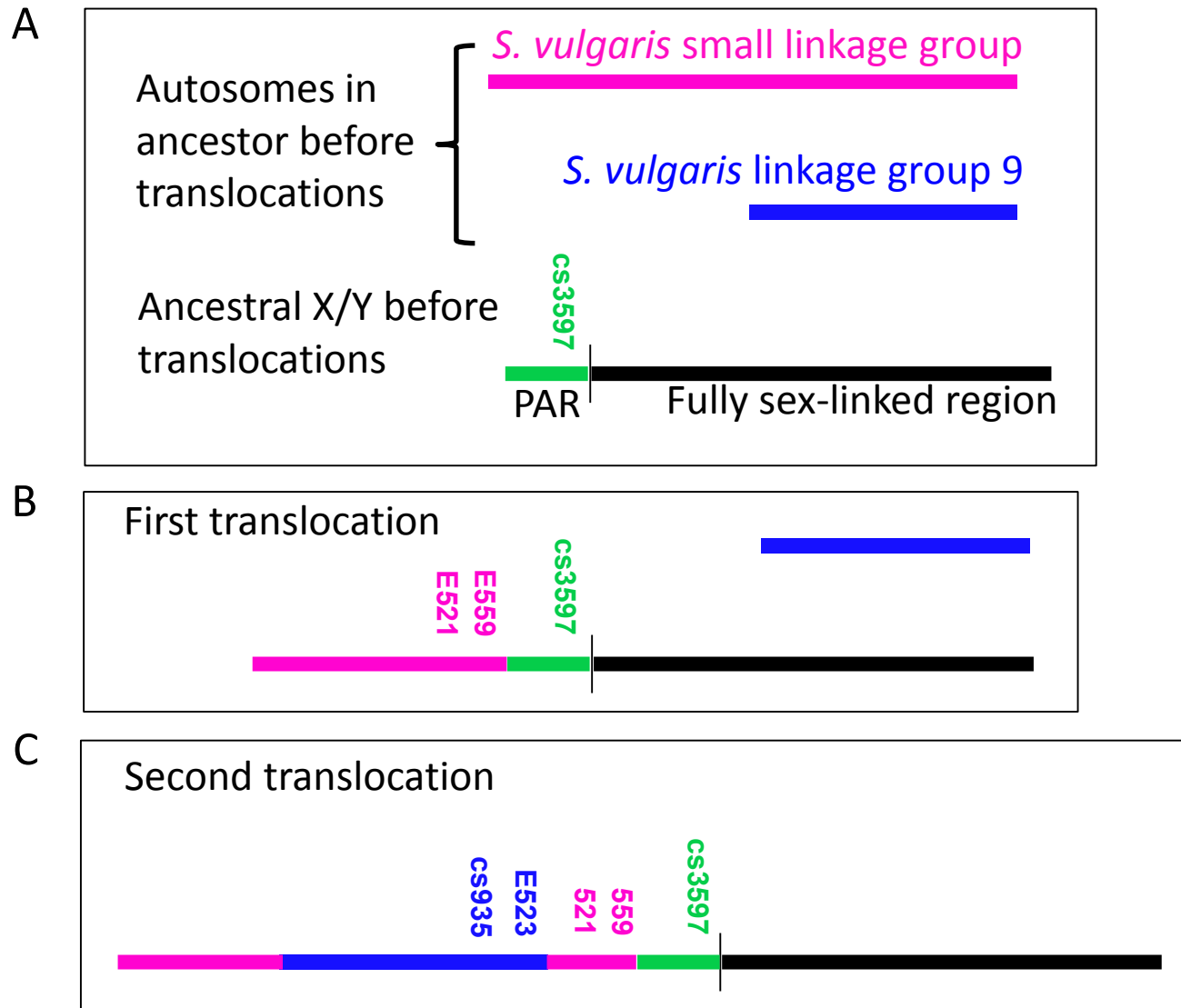


Figure 2

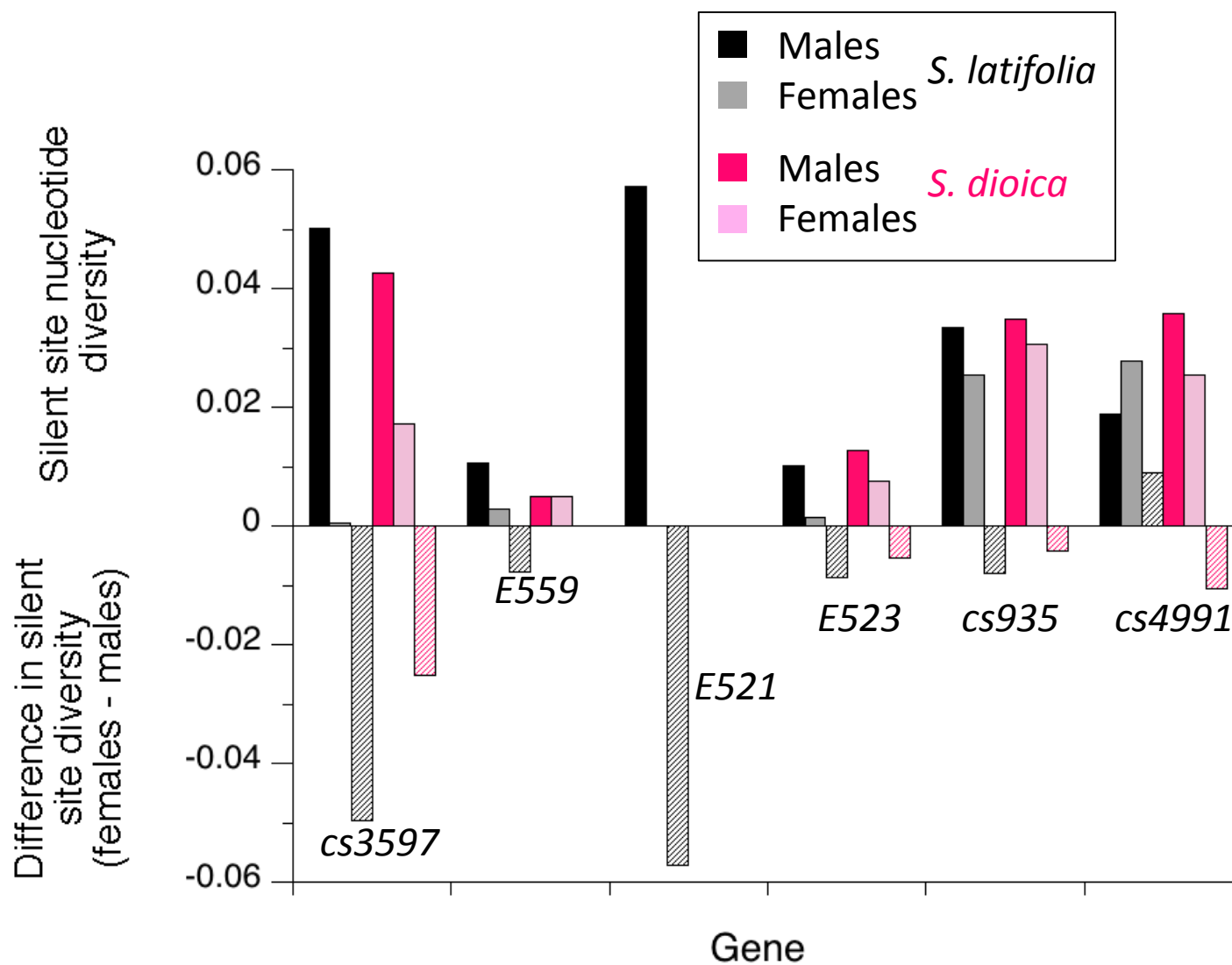
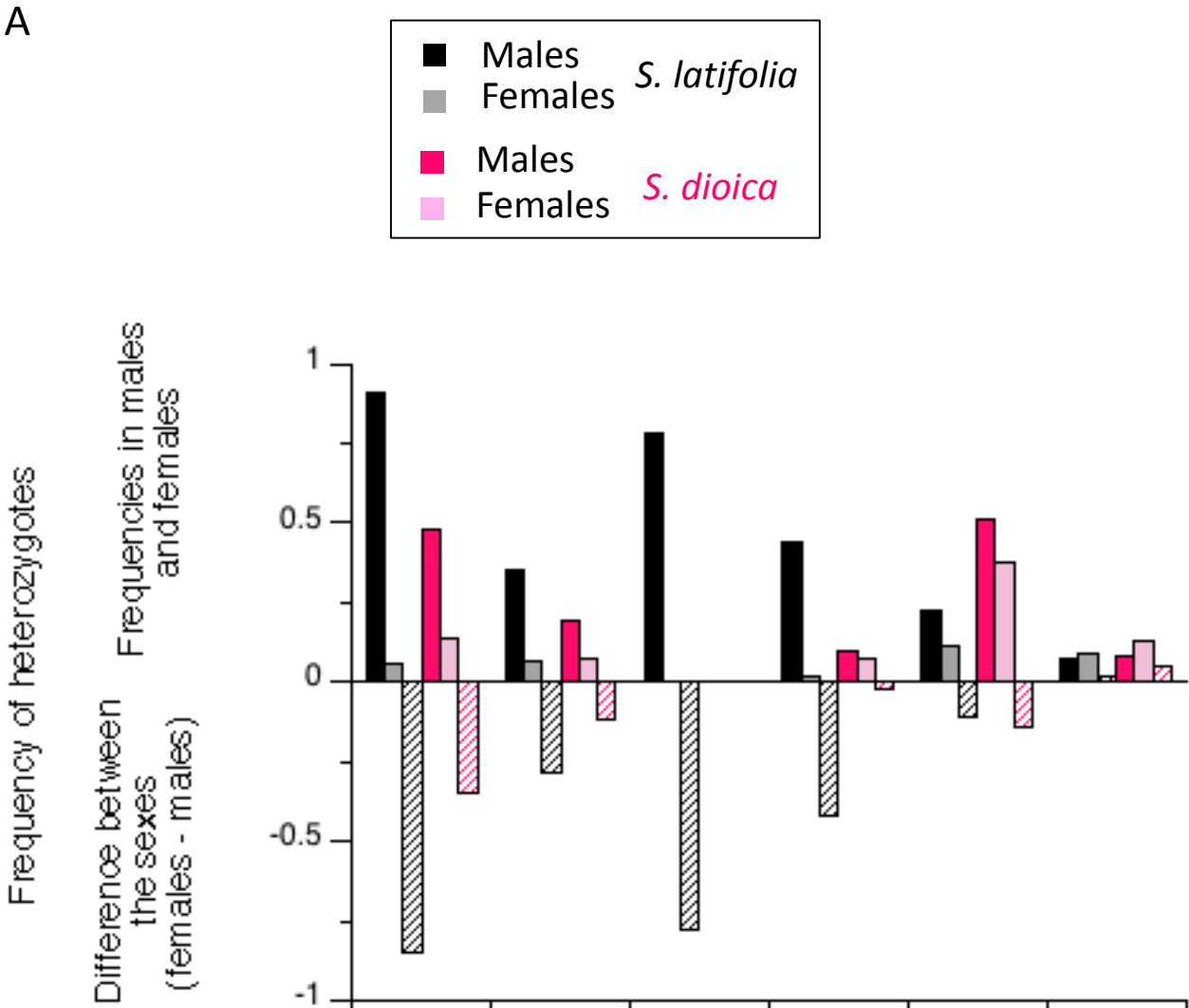


Figure 3

A



B

